

# Diagnostic – Problème de haute disponibilité dans mon cluster Proxmox

## Contexte de départ

- Cluster Proxmox à 3 nœuds : `Proxmox1`, `Proxmox2`, `Proxmox3`
- HA activée pour plusieurs VMs et conteneurs
- Watchdog (**mécanisme de sécurité matériel ou logiciel**, qui surveille si un système est encore vivant) matériel activé sur tous les nœuds
- Stockage partagé utilisé pour les VMs (Ceph)
- Le quorum est atteint en condition normale

## Comportement observé

1. Lorsqu'un nœud est arrêté proprement (`systemctl stop corosync`) ou alors avec un arrêt matériel propre :
  - **Les VMs/CT migrent et redémarrent correctement sur les autres nœuds.**
2. Lorsqu'un nœud **crashe brutalement** (coupure réseau, kernel panic, débranchement) :
  - **Les VMs sont bien migrées automatiquement par le gestionnaire HA.**
  - **Elles ne redémarrent pas correctement** (boot loop, plantage, écran noir).
  - **Impossible d'accéder à la console VM via l'interface Proxmox.**

- Impossible d'accéder aux consoles de mes différents noeuds.
- Les VMs restent **bloquées jusqu'au retour du nœud défaillant**, puis tout se débloque.

## Diagnostic

Le HA est bien configuré, **mais les redémarrages automatiques échouent suite à une bascule, ce qui provoque :**

- Des redémarrages en boucle ou plantages des VMs après migration
- Un comportement instable du cluster en situation de crash réel

## Vérifications effectuées

- `ha-manager status` OK
- `pvecm status` OK (quorum maintenu)
- Watchdog activé dans `/etc/default/pve-ha-manager`
- HA group configuré correctement
- Stockage partagé utilisé
- Les VMs migrent bien automatiquement

## Mon avis

Je pense que les problèmes que je rencontre avec la haute dispo viennent surtout du fait que **mes nœuds ne sont pas du tout uniforme au niveau matériel**. J'ai des différences au niveau de la RAM, des CPU, et sûrement des perfs disques, ce qui peut clairement poser souci quand une VM est déplacée automatiquement sur un autre nœud. En plus de ça, j'utilise **Ceph avec des disques de tailles différentes sur chaque serveur**, donc le cluster est un peu déséquilibré. Ça peut jouer sur les

perfs globales ou même sur la stabilité, surtout quand plusieurs VMs redémarrent en même temps. Du coup, je pense que ce mélange de matériel hétérogène + Ceph pas totalement uniforme explique pourquoi certaines VMs ne redémarrent pas bien après un crash de nœud.

## Résolution du problème

Lorsque j'ai testé la haute disponibilité avec Ceph, mes machines virtuelles refusaient de redémarrer sur un autre nœud après une panne. Après analyse, j'ai découvert que cela venait de la configuration de mon pool Ceph, qui était en réplication 3/1 (3 copies, 1 requise pour continuer). J'avais configuré ce mode pour voir combien de GB j'aurais pu économiser sur mon stockage. Ce mode autorise techniquement l'accès aux données avec un seul replica disponible, mais il ne garantit pas l'intégrité des écritures. Résultat : même si la VM est déplacée automatiquement sur un autre nœud par Proxmox, le disque RBD devient inaccessible ou instable, ce qui bloque le démarrage.

Pour garantir que les ressources HA redémarrent correctement après une panne, il faut configurer le pool en 3/2. Ce réglage offre une vraie tolérance à la perte d'un OSD (ou d'un nœud complet), tout en maintenant les disques RBD disponibles et cohérents pour Proxmox. Grâce à cela, les VM et conteneurs redémarrent sans erreur sur les autres nœuds du cluster.